EXHIBIT A

DANIEL J. MULLER, SBN 193396
dmuller@venturahersey.com
VENTURA HERSEY & MULLER, LLP
1506 Hamilton Avenue
San Jose, California 95125
Telephone: (408) 512-3022
Facsimile: (408) 512-3023

Attorneys for Plaintiffs and the Class

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA – SAN FRANCISCO DIVISION

MICHAEL CHABON, DAVID HENRY HWANG, MATTHEW KLAM, RACHEL LOUISE SNYDER, AND AYELET WALDMAN, individually and on behalf of all others similarly situated, Plaintiffs, v. META PLATFORMS, INC., a Delaware Corporation, Defendant.	Case No. **CLASS ACTION COMPLAINT** <u>CLASS ACTION</u> **JURY TRIAL DEMANDED**

CLASS ACTION COMPLAINT

Plaintiffs Michael Chabon, David Henry Hwang, Matthew Klam, Rachel Louise Snyder, and Ayelet Waldman ("Plaintiffs"), individually and on behalf of all others similarly situated, bring this class action against Defendant Meta Platforms, Inc. Plaintiffs' allege as follows based upon personal knowledge as to themselves and their own acts, and upon information and belief as to all other matters:

NATURE OF ACTION

1. This is a class action lawsuit brought by Plaintiffs on behalf of themselves and a Class of authors holding copyrights in their published works arising from Meta's clear infringement of their intellectual property.

2. Meta's LLaMA (Large Language Model Meta AI) is a set of large language models created and maintained by Meta Platforms, Inc. A *large language model* is an AI software program designed to produce convincingly natural texts outputs in response user prompts.

3. Rather than being programmed in the traditional manner, a large language model is "trained" by copying massive amounts of text and extracting expressive information from it. The body of text is referred to as the *training dataset.*

4. Accordingly, a large language model's output is therefore entirely and uniquely reliant on the material in its training dataset. Every time it assembles a text output, the model relies on the information it extracted from its training dataset. Therefore, the decisions about the textual information it includes in the training dataset are deliberate and important choices.

5. Plaintiffs and Class members are authors of books, screenplays, novels, and other written works. Plaintiffs and Class members possess copyrights for the books and written works they created and published. Plaintiffs and Class members did not consent to the use of their copyrighted books as training materials for LLaMA.

6. Nevertheless, their copyrighted protected works were copied and ingested as part of training LLaMA. Plaintiffs' copyrighted books appear in the dataset that Meta has admitted to using to train LLaMA.

7.   A large language model's responses to user prompts or queries are entirely and uniquely dependent on the text contained in its training dataset, necessarily processing and analyzing the information contained in its training dataset to generate responses.

JURISDICTION AND VENUE

8.   This Court has subject matter jurisdiction of this action pursuant to 28 U.S.C. § 1331 because this case arises under the Copyright Act (17 U.S.C. § 501) and the Digital Millennium Copyright Act (17 U.S.C. § 1202).

9.   This Court has personal jurisdiction over Defendants pursuant to 18 U.S.C. §§ 1965(b) & (d), because they maintain their principal places of business in, and are thus residents of, this judicial district, maintain minimum contacts with the United States, this judicial district, and this State, and they intentionally avail themselves of the laws of the United States and this state by conducting a substantial amount of business in California. For these same reasons, venue properly lies in this District pursuant to 28 U.S.C. §§ 1391(a), (b) and (c).

PARTIES

A.   Plaintiffs

10.   Plaintiff Michael Chabon ("Plaintiff Chabon") is a resident of California. Plaintiff Chabon is an author who owns registered copyrights in several works, including but not limited to, *The Mysteries of Pittsburgh*, *Wonder Boys, The Amazing Adventures of Kavalier & Clay, the Yiddish Policemen's Union, Gentlemen of the Road, Telegraph Avenue,* and *Moonglow.* Plaintiff Chabon is the recipient of the Pulitzer Prize for Fiction, Hugo, Nebula, Los Angeles Times Book Prize, and the National Jewish Book Award, among many other awards received during the span of a writing career of more than 30 years. Plaintiff Chabon's works include copyright-management information that provides information about the copyrighted work, including the title of the work, its ISBN or copyright registration number, the name of the author, and the year of publication.

11.   Plaintiff David Henry Hwang ("Plaintiff Hwang") is a resident of New York. Plaintiff Hwang is a playwright and screenwriter who owns registered copyrights in several plays, including but not limited to, *M. Butterfly, Chinglish, Yellow Face, Golden Child, the*

1

Dance and the Railroad, and *FOB*, as well as the Broadway musicals *Aida, Flower Drum Song* (2002 revival) and *Disney's Tarzan*. Plaintiff Hwang is a Tony Award winner and three-time nominee, a Grammy Award winner and two time nominee, a three-time OBIE Award winner, and a three-time finalist for the Pulitzer Prize in Drama. Plaintiff Hwang's works include copyright-management information that provides information about the copyrighted work, including the title of the work, its ISBN or copyright registration number, the name of the author, and the year of publication.

12.   Plaintiff Matthew Klam ("Plaintiff Klam") is a resident of Washington D.C. Plaintiff Klam is an author who owns registered copyrights in several works, including but not limited to, *Who is Rich?,* and *Sam the Cat and Other Stories*. Plaintiff Klam is a recipient of a Guggenheim Fellowship, a Robert Bingham/PEN Award, a Whiting Writer's Award, and a National Endowment of the Arts. Plaintiff Klam's works have been selected as Notable Books of the year by *The New York Times, The Los Angeles Times, the Kansas City Star, and the Washington Post.* His work has appeared in *The New York Times, The New Yorker, Harper's Magazine*, and elsewhere. Plaintiff Klam's works include copyright-management information that provides information about the copyrighted work, including the title of the work, its ISBN or copyright registration number, the name of the author, and the year of publication.

13.   Plaintiff Rachel Louise Snyder ("Plaintiff Snyder") is a resident of Washington, D.C. Plaintiff Snyder is an author who owns registered copyrights in several works, including but not limited to, *Women We Buried, Women We Burned*, *No Visible Bruises – What We Don't Know About Domestic Violence Can Kill Us*, *What We've Lost is Nothing*, and *Fugitive Denim: A Moving Story of People and Pants in the Borderless World of Global Trade*. Plaintiff Snyder is the recipient of the J. Anthony Lukas Work-in-Progress Award, the Hillman Prize, and the Helen Bernstein Book Award, and finalist for the National Book Critics Circle Award, *Los Angeles Times* Book Prize, and Kirkus Award. Her work has appeared in *The New Yorker*, *The New York Times*, *Slate*, and elsewhere. Plaintiff Snyder's works include copyright-management information that provides information about the copyrighted work, including the

CLASS ACTION COMPLAINT

title of the work, its ISBN or copyright registration number, the name of the author, and the year of publication.

14.   Plaintiff Ayelet Waldman ("Plaintiff Waldman") is a resident of California. Plaintiff Waldman is an author and screen and television writer who owns registered copyrights in several works, including but not limited to, *Love and other Impossible Pursuits*, *Red Hook Road, Love and Treasure*, *Bad Mother, Daughter's Keeper, A Really Good Day,* and *Mommy Track Mysteries*. Plaintiff Waldman has been nominated for an Emmy and Golden Globe and is the recipient of numerous awards including a Peabody, AFI award, and a Pen Award, among others. Plaintiff Waldman's works include copyright-management information that provides information about the copyrighted work, including the title of the work, its ISBN or copyright registration number, the name of the author, and the year of publication.

15.   At all times relevant hereto, Plaintiffs have been and remain the holders of the exclusive rights under the Copyright Act of 1976 (17 U.S.C. §§ 101, *et seq.* and all amendments thereto) to reproduce, distribute, display, or license the reproduction, distribution, and/or display the works identified in paragraphs 13-17, *supra*.

B.   Defendant

16.   Defendant Meta is a Delaware corporation with its principal place of business at 1601 Willow Road, Menlo Park, California 94025.

AGENTS AND CO-CONSPIRATORS

17.   The unlawful acts alleged against Meta in this class action complaint were authorized, ordered, or performed by the Defendant's respective officers, agents, employees, representatives, or shareholders while actively engaged in the management, direction, or control of the Defendant's businesses or affairs. The Defendant's agents operated under the explicit and apparent authority of their principals. Each Defendant, and its subsidiaries, affiliates, and agents operated as a single unified entity.

18.   Various persons and/or firms not named as Defendants may have participated as co-conspirators in the violations alleged herein and may have performed acts and made

CLASS ACTION COMPLAINT

statements in furtherance thereof. Each acted as the principal, agent or joint venture of, or for other Defendants with respect to the acts, violations, and common course of conduct alleged herein.

FACTUAL ALLEGATIONS

A.   Meta Platform's Artificial Intelligence Products

19.   Meta creates, markets, and sells software and hardware technology products, including Facebook, Instagram, and Horizon Worlds. Meta also has a large artificial-intelligence group called Meta AI that creates and distributes artificial-intelligence software products.

20.   AI software is designed to algorithmically simulate human reasoning or inference, often based upon statistical models or methods.

21.   In February 2023, Meta released an AI product called LLaMA. LLaMA is a set of large language models. A large language model (or "LLM" for short) is AI software designed to parse and emit natural language. Though a large language model is a software program, it is not created the way most software programs are—that is, by human software engineers writing code. Rather, a large language model is "trained" by copying massive amounts of text from various sources and feeding these copies into the model. This corpus of input material is called the training dataset. During training, the large language model copies each piece of text in the training dataset and extracts expressive information from it. The large language model progressively adjusts its output to more closely resemble the sequences of words copied from the training dataset. Once the large language model copies and ingests the all of this text, it is able to generate and produce convincing simulations of natural written language as it appears in the training dataset.

22.   Much of the material in Meta's training dataset, however, comes from copyrighted works—including works written by Plaintiffs—that were copied by Meta without consent, without credit, and without compensation.

23.   Plaintiffs published written works contain certain copyright management information. This information includes the written work's title, the ISBN number or copyright number, the author's name the copyright holder's name, and terms and conditions of use.

CLASS ACTION COMPLAINT

24. Meta introduced LLaMA in a paper called "LLaMA: Open and Efficient Foundation Language Models". In the paper, Meta describes the LLaMA training dataset as "a large quantity of textual data" that was chosen because it was "publicly available, and compatible with open sourcing."

25. *Open sourcing* refers to putting data under a permissive style of copyright license called an *open-source license*. Copyrighted materials, however, are not ordinarily "compatible with open sourcing" unless and until the copyright owner first places the material under an open-source license, thereby enabling others to do so later.

26. In a table describing the composition of the LLaMA training dataset, Meta notes that 85 gigabytes of the training data comes from a category called "Books." Meta further elaborates that "Books" comprises the text of books from two internet sources: (1) Project Gutenberg, an online archive of approximately 70,000 books that are out of copyright, and (2) "the Books3 section of ThePile . . . a publicly available dataset for training large language models." Meta's paper on LLaMA does not further describe the contents of Books3 or ThePile.

27. In a table describing the composition of the LLaMA training dataset, Meta notes that 85 gigabytes of the training data comes from a category called "Books." Meta further elaborates that "Books" comprises the text of books from two internet sources: (1) Project Gutenberg, an online archive of approximately 70,000 books that are out of copyright, and (2) "the Books3 section of ThePile . . . a publicly available dataset for training large language models." Meta's paper on LLaMA does not further describe the contents of Books3 or ThePile.

28. But that information is available elsewhere. ThePile is a dataset assembled by a research organization called EleutherAI. In December 2020, EleutherAI introduced this dataset in a paper called "The Pile: An 800GB Dataset of Diverse Text for Language Modeling".

29. The EleutherAI paper reveals that the Books3 dataset comprises 108 gigabytes of data, or approximately 12% of the dataset, making it the third largest component of The Pile by size.

30. The EleutherAI paper describes the contents of Books3:

Books3 is a dataset of books derived from a copy of the contents of the

CLASS ACTION COMPLAINT

> Bibliotik private tracker … Bibliotik consists of a mix of fiction and nonfiction books and is almost an order of magnitude larger than our next largest book dataset (BookCorpus2). We included Bibliotik because books are invaluable for long-range context modeling research and coherent storytelling.

31. Bibliotik is one of a number of notorious "shadow library" websites that also includes Library Genesis (aka LibGen), Z-Library (aka B-ok), and Sci-Hub. The books and other materials aggregated by these websites have also been available in bulk via torrent systems. These shadow libraries have long been of interest to the AI-training community because of the large quantity of copyrighted material they contain. For that reason, these shadow libraries are also flagrantly illegal.

32. The person who assembled the Books3 dataset has confirmed in public statements that it represents "all of Bibliotik" and contains 196,640 books. EleutherAI currently distributes copies of Books3 through its website (https://pile.eleuther.ai/).

33. The Books3 dataset is also available from a popular AI project hosting service called Hugging Face (https://huggingface.co/datasets/the_pile_books3).

34. Many of Plaintiffs' written works appear in the Books3 dataset, these written works are referred to as the Infringed Works.

35. For example, Books3 contains a significant amount of Plaintiff Chabon's works, including, but not limited to, *The Final Solution, Bookends: Collected Intros and Outros, Kingdom of Olives and Ash, Manhood for Amateurs: The Pleasures and Regrets of a Husband, Father, and Son, Maps and Legends, McSweeney's Mammoth Treasury of Thrilling Tales, Werewolves in Their Youth, Michael Chabon's America: Magical Words, Secret Worlds, and Sacred Spaces, Moonglow, Pops Fatherhood in Pieces, The Amazing Adventures of Kavalier & Clay, and the Yiddish Policemen's Union*.

36. Books3 similarly contains Plaintiff Hwang's written works, including, but not limited to, *Golden Child, M. Butterfly,* and *Trying to Find Chinatown*.

37. Plaintiff Klam's works are similarly found in the Books3 dataset, including, but not limited to, *Who is Rich?* and *Sam the Cat*.

CLASS ACTION COMPLAINT

38. Plaintiff Snyder's works also are contained in the Books3 dataset, including, but not limited to, *No Visible Bruises: What We Don'y Know about Domestic Violence Can Kill Us.*

39. In the same vein, Plaintiff Waldman's works appear in the Books3 dataset, including, but not limited to, *A Really Good Day, Bad Mother, Love and Other Impossible Pursuits, and Love and Treasure.*

40. Since the launch of the LLaMA language models in February 2023, Meta has made these models selectively available to organizations that request access, saying:

> To maintain integrity and prevent misuse, we are releasing our model under a noncommercial license focused on research use cases. Access to the model will be granted on a case-by-case basis to academic researchers; those affiliated with organizations in government, civil society, and academia; and industry research laboratories around the world. People interested in applying for access can find the link to the application in our research paper.

41. Meta has not disclosed what criteria it uses to decide who is eligible to receive the LLaMA language models, nor who has actually received them, or whether Meta has in fact adhered to its stated criteria. On information and belief, Meta has in fact distributed the LLaMA models to certain people and entities, continues to do so, and has benefited financially from these acts.

42. In March 2023, the LLaMA language models were leaked to a public internet site and have continued to circulate. Meta has not disclosed what role it had, if any, in the leak.

43. Later in March 2023, Meta issued a DMCA takedown notice to a programmer on GitHub who had released a tool that helped users download the leaked LLaMA language models. In the notice, Meta asserted copyright over the LLaMA language models.

44. According to reporting in June 2023, Meta plans to make the next version of LLaMA commercially available.

CLASS ALLEGATIONS

45. Plaintiffs bring this action pursuant to the provisions of Rules 23(a), 23(b)(2), and 23(b)(3) of the Federal Rules of Civil Procedure, on behalf of themselves and the following proposed Class:

7

CLASS ACTION COMPLAINT

All persons or entities in the United States that own a United States copyright in any work that was used as training data for the LLaMA language models during the Class Period.

44.   Excluded from the Class are Defendant, its employees, officers, directors, legal representatives, heirs, successors, wholly- or partly-owned, and its subsidiaries and affiliates; proposed Class counsel and their employees; the judicial officers and associated court staff assigned to this case and their immediate family members; all persons who make a timely election to be excluded from the Class; governmental entities; and the judge to whom this case is assigned and his/her immediate family.

45.   This action has been brought and may be properly maintained on behalf of the Class proposed herein under Federal Rule of Civil Procedure 23.

46.   Numerosity. Federal Rule of Civil Procedure 23(a)(1): The members of the Class are so numerous and geographically dispersed that individual joinder of all Class members is impracticable. On information and belief, there are at least tens of thousands of members in the Class. The Class members may be easily derived from Defendants' records.

47.   Commonality and Predominance. Federal Rule of Civil Procedure 23(a)(2) and 23(b)(3): This action involves common questions of law and fact, which predominate over any questions affecting individual Class members, including, without limitation:

    a.   Whether Defendant violated the copyrights of Plaintiffs and the Class when they downloaded copies of Plaintiffs' and the Class's Infringed Works and used them to train the LLaMA language models;

    b.   Whether the LLaMA language models are themselves infringing derivative works based on Plaintiffs' and the Class's Infringed Works;

    c.   Whether the text outputs of the LLaMA language models are infringing derivative works based on Plaintiffs' Infringed Works;

    d.   Whether Defendant violate the DMCA by removing copyright-management information from Plaintiffs' and the Class's Infringed Works;

    e.   Whether Defendant was unjustly enriched by the unlawful conduct alleged herein;

8

CLASS ACTION COMPLAINT

f. Whether Defendant's conduct allege herein constitutes Unfair Competition under California Business and Professions Code Secion 17200 *et seq.*

g. Whether Defendant's conduct alleged herein constitute common unfair competition;

h. Whether any affirmative defense excuses Defendant's conduct;

i. Whether any statutes of limitation limits Plaintiffs' and the Class's potential for recovery;

j. Whether Plaintiffs and the other Class members are entitled to equitable relief, including, but not limited to, restitution or injunctive relief; and

k. Whether Plaintiffs and the other Class members are entitled to damages and other monetary relief and, if so, in what amount.

48. <u>Typicality</u>. Federal Rule of Civil Procedure 23(a)(3): Plaintiffs' claims are typical of the other Class members' claims because, among other things, all Class members were comparably injured through Defendant's wrongful conduct as described above.

49. <u>Adequacy</u>. Federal Rule of Civil Procedure 23(a)(4): Plaintiffs are adequate Class representative because their interests do not conflict with the interests of the other members of the Class they seeks to represent; Plaintiffs have retained counsel competent and experienced in complex class action litigation; and Plaintiffs intend to prosecute this action vigorously. The interests of the Class will be fairly and adequately protected by Plaintiffs and their counsel.

50. <u>Declaratory and Injunctive Relief</u>. Federal Rule of Civil Procedure 23(b)(2): Defendants have acted or refused to act on grounds generally applicable to Plaintiffs and the other members of the Class, thereby making appropriate final injunctive relief and declaratory relief with respect to the Class as a whole.

51. <u>Superiority</u>. Federal Rule of Civil Procedure 23(b)(3): A class action is superior to any other available means for the fair and efficient adjudication of this controversy, and no unusual difficulties are likely to be encountered in the management of this class action. The damages or other financial detriment suffered by Plaintiffs and the other Class members are

9

relatively small compared to the burden and expense that would be required to individually litigate their claims against Defendants, so it would be impracticable for the members of the Class to individually seek redress for Defendant's wrongful conduct. Even if Class members could afford individual litigation, the court system could not. Individualized litigation creates a potential for inconsistent or contradictory judgments, and increases the delay and expense to all parties and the court system. By contrast, the class action device presents far fewer management difficulties, and provides the benefits of single adjudication, economy of scale, and comprehensive supervision by a single court.

CAUSES OF ACTION

FIRST CAUSE OF ACTION

DIRECT COPYRIGHT INFRINGEMENT,
17 U.S.C. § 106, *et seq.*

52.   Plaintiffs hereby incorporate by reference the allegations contained in the preceding paragraphs of this Complaint.

53.   Plaintiffs bring this claim on behalf of themselves and on behalf of the Class against Defendants.

54.   As the owners of the registered copyrights in the Infringed Works, Plaintiffs hold the exclusive rights to those books under 17 U.S.C. § 106.

55.   To train the LLaMA language models, Meta copied the Books3 dataset, which includes the Infringed Works.

56.   Plaintiffs never authorized Meta to make copies of their Infringed Works, make derivative works, publicly display copies (or derivative works), or distribute copies (or derivative works). All those rights belong exclusively to Plaintiffs under copyright law.

57.   Meta made copies of the Infringed Works during the training process of the LLaMA language models without Plaintiffs' permission.

58.   Because the LLaMA language models cannot function without the expressive information extracted from Infringed Works and retained inside the LLaMA language models,

these LLaMA language models are themselves infringing derivative works, made without Plaintiffs' permission and in violation of their exclusive rights under the Copyright Act.

59.   Plaintiffs and the Class have been injured by Meta's acts of direct copyright infringement. Plaintiffs and the Class are entitled to statutory damages, actual damages, restitution of profits, and other remedies provided by law.

SECOND CAUSE OF ACTION

VICARIOUS COPYRIGHT INFRINGEMENT
17 U.S.C. § 106

60.   Plaintiffs incorporate by reference all allegations of the preceding paragraphs as though fully set forth herein.

61.   Plaintiffs bring this claim on behalf of herself and on behalf of the Class against Defendants.

62.   Because the output of the LLaMA language models is based on expressive information extracted from Plaintiffs' Infringed Works, every output of the LLaMA language models is an infringing derivative work, made without Plaintiffs' permission and in violation of their exclusive rights under the Copyright Act.

63.   Meta has the right and ability to control the output of the LLaMA language models. Meta has benefited financially from the infringing output of the LLaMA language models. Therefore, every output from the LLaMA language models constitutes an act of vicarious copyright infringement.

64.   Plaintiffs and the Class have been injured by Meta's acts of vicarious copyright infringement. Plaintiffs and the Class are entitled to statutory damages, actual damages, restitution of profits, and other remedies provided by law.

THIRD CAUSE OF ACTION

DIGITAL MILLENNIUM COPYRIGHT ACT – REMOVAL OF COPYRIGHT
MANAGEMENT INFORMATION
17 U.S.C. § 1202(B)

65.   Plaintiffs incorporate by reference all allegations of the preceding paragraphs as

though fully set forth herein.

66.   Plaintiffs bring this claim on behalf of herself and on behalf of the Class against Defendants.

67.   Plaintiffs included one or more forms of copyright-management information ("CMI") in each of the Infringed Works, including: copyright notice, title and other identifying information, or the name or other identifying information about the owners of each book, terms and conditions of use, and identifying numbers or symbols referring to CMI.

68.   Without the authority of Plaintiffs and the Class, Meta copied the Infringed Works and used them as training data for the LLaMA language models. By design, the training process does not preserve any CMI. Therefore, Meta intentionally removed CMI from the Infringed Works in violation of 17 U.S.C. § 1202(b)(1).

69.   Without the authority of Plaintiffs and the Class, Defendant created derivative works based on the Infringed Works. By distributing these works without their CMI, Meta violated 17 U.S.C. § 1202(b)(3).

70.   By falsely claiming that it has sole copyright in the LLaMA language models—which it cannot, because the LLaMA language models are infringing derivative works—Meta violated 17 U.S.C. § 1202(a)(1).

71.   Meta knew or had reasonable grounds to know that this removal of CMI would facilitate copyright infringement by concealing the fact that every output from the LLaMA language models is an infringing derivative work, synthesized entirely from expressive information found in the training data.

72.   Plaintiffs and the Class have been injured by Meta's removal of CMI. Plaintiffs and the Class are entitled to statutory damages, actual damages, restitution of profits, and other remedies provided by law.

FOURTH CAUSE OF ACTION

VIOLATIONS OF THE CALIFORNIA UNFAIR COMPETITION LAW
CAL. BUS. & PROF. CODE §§ 17200, *ET SEQ.*

73.   Plaintiffs and the Class incorporate by reference each preceding and succeeding paragraph as though fully set forth at length herein.

74.   Plaintiffs bring this claim on behalf of herself and on behalf of the Class against Defendants.

75.   Defendant has engaged in unlawful business practices, including violating Plaintiffs' and the Class's rights under the DMCA, and using the Infringed Works to train LLaMA without Plaintiffs' or the Class's authorization.

76.   The unlawful business practices described herein violate California Business and Professions Code section 17200 et seq. because that conduct is otherwise unlawful by violating the DMCA.

77.   The unlawful business practices described herein violate California Business and Professions Code section 17200 et seq. because they are unfair, immoral, unethical, oppressive, unscrupulous or injurious to consumers, because, among other reasons, Defendant used Plaintiffs' protected works to train LLaMA for Defendant's own gain without Plaintiffs' and the Class's authorization.

78.   The unlawful business practices described herein violate California Business and Professions Code section 17200 et seq. as fraudulent because consumers are likely to be deceived because, among other reasons, Meta caused LLaMA's output to be emitted without any credit to Plaintiffs' or the Class whose Infringed Works comprise LLaMA's training dataset.

79.   Plaintiffs and the Class have been injured by Meta's removal of CMI. Plaintiffs and the Class are entitled to statutory damages, actual damages, restitution of profits, and other remedies provided by law.

FIFTH CAUSE OF ACTION

NEGLIGENCE

80.   Plaintiffs incorporate by reference the allegations of all foregoing paragraphs as

13

1   if they had been set forth in full herein.

2   81.   Plaintiffs bring this claim on behalf of themselves and on behalf of the Class

3   against Defendants.

4   82.   Defendant owed a duty of care toward Plaintiffs and the Class based upon

5   Defendant's relationship to them. This duty is based upon Defendant's obligations, custom and

6   practice, right to control information in its possession, exercise of control over the information

7   in its possession, authority to control the information in its possession, and the commission of

8   affirmative acts that result in said harms and losses. Additionally, this duty is based on the

9   requirements of California Civil Code section 1714, requiring all "persons," including

10   Defendant, to act in a reasonable manner toward others.

11   83.   Defendant breached its duties by negligently, carelessly, and recklessly

12   collecting, maintaining and controlling Plaintiffs' and Class members' Infringed Works and

13   engineering, designing, maintaining and controlling systems—including LLaMA—which are

14   trained on Plaintiffs' and Class members' Infringed Works without their authorization.

15   84.   Defendant owed Plaintiffs and Class members a duty of care to maintain the

16   Infringed Works once collected and ingested for training LLaMA.

17   85.   Defendant also owed Plaintiffs and Class members a duty of care to not use the

18   Infringed Works in a way that would foreseeably cause Plaintiffs and Class members injury, for

19   instance, by using the Infringed Works to train LLaMA.

20   86.   Defendant breached their duties by, inter alia, the Infringed Works to train

21   LLaMA.

22   **SIXTH CAUSE OF ACTION**

23   **UNJUST ENRICHMENT**

24   87.   Plaintiffs incorporate by reference all allegations of the preceding paragraphs as

25   though fully set forth herein.

26   88.   Plaintiffs and the Class have invested substantial time and energy in creating the

27   Infringed Works.

28

14

CLASS ACTION COMPLAINT

89. Defendants have unjustly utilized access to the Infringed Materials to train LLaMA.

90. Plaintiffs did not consent to the unauthorized use of the Infringed Materials to train LLaMA.

91. By using Plaintiffs' Infringed Works to train LLaMA, Plaintiffs and the Class were deprived of the benefits of their work, including monetary damages.

92. Defendants derived or intend to derive profit and other benefits from the use of the Infringed Materials to train LLaMA.

93. It would be unjust for Defendant to retain those benefits.

94. The conduct of Defendant is causing and, unless enjoined and restrained by this Court, will continue to cause Plaintiffs and the Class great and irreparable injury that cannot fully be compensated or measured in money.

REQUEST FOR RELIEF

WHEREFORE, Plaintiffs, individually and on behalf of members of the Class defined above, respectfully request that the Court enter judgment against Defendants and award the following relief:

A. Certification of this action as a class action pursuant to Rule 23 of the Federal Rules of Civil Procedure, declaring Plaintiffs as the representative of the Class, and Plaintiffs' counsel as counsel for the Class;

B. An order awarding declaratory relief and temporarily and permanently enjoining Defendant from continuing the unlawful and unfair business practices alleged in this Complaint and to ensure that all applicable information set forth in 17 U.S.C. § 1203(b)(1) is included when appropriate;

C. An award of statutory and other damages under 17 U.S.C. § 504 for violations of the copyrights of Plaintiff and the Class by Defendants.

D. An award of statutory damages under 17 U.S.C. § 1203(b)(3) and 17 U.S.C. § 1203(c)(3), or in the alternative, an award of actual damages and any additional profits under 17 U.S.C. § 1203(c)(2);

15

CLASS ACTION COMPLAINT

1   E.   A declaration that Defendant is financially responsible for all Class notice and

2 the administration of Class relief;

3   F.   An order awarding any applicable statutory and civil penalties;

4   G.   An order requiring Defendant to pay both pre- and post-judgment interest on any

5 amounts awarded;

6   H.   An award of costs, expenses, and attorneys' fees as permitted by law; and

7   I.   Such other or further relief as the Court may deem appropriate, just, and

8 equitable.

9 <div align="center">**DEMAND FOR JURY TRIAL**</div>

10   Plaintiffs hereby demand a jury trial for all claims so triable.

11 DATED: September 12, 2023   Respectfully submitted,

12

13   /s/ *Daniel J. Muller*
DANIEL J. MULLER, SBN 193396

14 dmuller@venturahersey.com
VENTURA HERSEY & MULLER, LLP

15 1506 Hamilton Avenue
San Jose, California 95125

16 Telephone: (408) 512-3022
Facsimile: (408) 512-3023

17 dmuller@venturahersey.com

18

19   /s/ *Bryan L. Clobes*
Bryan L. Clobes (*pro hac vice anticipated*)

20 **CAFFERTY CLOBES MERIWETHER
    & SPRENGEL LLP**

21 205 N. Monroe Street
Media, PA 19063

22 Tel: 215-864-2800
bclobes@caffertyclobes.com

23

24 Alexander J. Sweatman (*pro hac vice anticipated*)
**CAFFERTY CLOBES MERIWETHER
    & SPRENGEL LLP**

25 135 South LaSalle Street, Suite 3210

26 Chicago, IL 60603
Tel: 312-782-4880

27 asweatman@caffertyclobes.com
Attorneys for Plaintiffs

28

<div align="center">16</div>
<div align="center">CLASS ACTION COMPLAINT</div>